5633?;
N90-22950

A COMMENTARY ON PERCEPTION–ACTION RELATIONSHIPS IN SPATIAL DISPLAY INSTRUMENTS

Wayne L. Shebilske
Department of Psychology
Texas A&M University
College Station, Texas

SUMMARY

My presentation at the conference was based on a paper that was prepared in advance and submitted for publication in this volume. In addition, the presentation included several ideas that emerged during the conference as a result of interactions with other participants. I would like to convey those ideas here along with other thoughts that occurred to me later. I will organize this commentary around three objectives: (1) to promote transfer of information across disciplines; (2) to caution basic and applied researchers about the danger of assuming simple relationships between stimulus information, perceptual impressions, and performance including pattern recognition and sensorimotor skills; and (3) to develop a theoretical and empirical foundation for predicting those relationships.

INFORMATION TRANSFER ACROSS DISCIPLINES

This conference clearly indicated that basic and applied researchers have crossed traditional boundaries to work together toward new applications of spatial display instruments. For example, on the one hand, leaders in basic research on perception, such as Richard Gregory and Richard Held, spoke about their current research concerning applications of spatial display instruments. On the other hand, M. W. McGreevy, a leader in promoting the application of spatial display in space, also promoted basic research on sensation and perception. Thus, in place of the bottlenecks of which I spoke in my paper, I got an impression of open communication and a steady flow of information. As a result, multidisciplinary research teams have exciting agendas for research on general principles that have direct relevance to spatial display technology.

I also discovered tremendous interest in transferring information between those who are developing spatial display instruments to enhance normal sensory function or to extend it to remote-control situations and those who are developing electronic aids for the blind. I discussed with many participants of the present conference a study on the latter topic that was organized while I was Study Director for the committee on Vision (COVIS) of the National Academy of Sciences. That Committee has recently released a study on electronic aids for the blind that includes a research agenda that is highly relevant to the research programs of many of those who participated in the present conference. For example, the report calls for more research on the nature of information that is picked up about surfaces, and we saw in the present conference that this issue is also important in teleoperation of land vehicles (see McGovern, this volume). The COVIS report can be ordered by calling (202) 334-2565. You might also want to request information on a recent COVIS conference on visual displays.

DANGERS OF ASSUMING SIMPLE RELATIONSHIPS BETWEEN PERCEPTUAL IMPRESSIONS AND PERFORMANCE: DON'T TRUST YOUR INTUITIONS

My paper reviewed evidence that relationships between stimulus information, perceptual impressions, and performance is complex, variable, and currently unpredictable. It is tempting to treat the evidence as quirks since that would make life so much easier for basic and applied researchers. If these relationships were simple, constant, and predictable, consider how worry-free one could be in making inferences about basic principles of perception from observations about performance, or in making decisions about human factors of performance from data about perceptual impressions. Several considerations add to the temptation to regard the evidence as quirks. For one thing, much of it comes from exotic clinical or laboratory situations regarding blind sight, subliminal priming of recognition, and paradoxical perceptions. Furthermore, our intuitions tell us that our sensory-guided performance corresponds to our perceptions most of the time.

With these considerations in mind, my presentation included a simple demonstration of discordance between perceptual impressions and performance in an everyday situation. I placed a plastic golf ball on a carry-out lid on an old McDonald's coffee cup and asked people to observe the ball with one or two eyes. The ball and cup were placed on an edge of a table while observers stood leaning over the cup and judged the apparent viewing distance between themselves and the ball. In agreement with data reviewed by Stanley Roscoe (this volume), participants at the conference and undergraduates tested at Texas A&M University saw the ball as being the same distance or slightly farther away (an average of about 1 cm) with one eye in comparison to the apparent distance with binocular viewing. The same observers were also asked to hit the ball off the cup by swinging a ruler parallel to the cup surface at the level of the ball. Order of these tasks was counterbalanced across subjects and the results were the same both groups. Almost all subjects swung well above the ball (an average of about 3 cm). I call the results of this demonstration the Old McDonald effect. The demonstration is easy to repeat. You may substitute a Coke can, or any other small can, and a wadded piece of paper for the coffee cup and ball. You may also try to hit the paper with your finger instead of a ruler, as long as you attempt to make one smooth, rapid swing parallel to the surface of the stand. If you are among the many people who are surprised to see themselves swing above the ball, you will be in a better position to appreciate the point of the demonstration, which is that *you cannot trust your intuitions about perceptual impressions and performance, even in over-learned skills such as hitting objects with your hand in natural conditions*. This is the main take-home message that I tried to emphasize in my presentation.

This message is relevant to other projects that were presented at the conference. For example, some simulators have displays that are so realistic that an observer gets an impression of actually being at the scene that is displayed, and scientists are attempting to analyze the determinants of telepresence (see Held, this volume). Held outlined a framework for analyzing determinants of the compellingness of these impressions, including time lags in visuo-motor tasks. The distinction between perceptual impressions and performance will be critical in this context if it turns out that the factors influencing perceptual compellingness are different than those determining proficiency of performance. Similarly, those who are studying stereopsis (e.g., Enright, this volume; Foley, this volume; Schor, this volume; and Stevens, this volume) might find different factors affecting impressions of depth and performance with 3D displays. Finally, efforts are being made to train pilots to see relative vertical separations better in collision-avoidance situations (Sherry Chappell, personal communications, September 1, 1987). Scientists might also find here that factors influence perceptual impressions and performance differently.

My short-term goal is to alert applied and basic researchers about potential discrepancies between the determinants of perceptual impressions and performance that could affect their research and instrumentation designs. For now, scientists will have to watch their step on a case-by-case basis since there are no empirically founded principles that would enable general predictions. The last section of this commentary will turn to my long-term goal of providing a foundation for such predictions.

THEORETICAL AND EMPIRICAL FOUNDATIONS FOR PREDICTING RELATIONSHIPS BETWEEN STIMULUS INFORMATION, PERCEPTUAL IMPRESSIONS, AND PERFORMANCE

This volume contains three hypotheses that propose a framework within which to investigate the many-to-one relationship that exists between stimulus information, perceptual impressions, and performance: (1) the Perception Plus Transformation hypothesis (see Foley, this volume); (2) the Dual Mode of Visual Representation hypothesis (see Bridgeman, this volume); and (3) the Ecologically Insulated Event Input Operations (EIEIO) hypothesis. Figure 1 illustrates all three. They all begin with conversion of distal information, which is in the environment, into proximal information, which is at the interface between the environment and the sensory system. According to the Perception Plus Transformation hypothesis, proximal information is converted into abstract symbolic representations that result in perceptions and sensory-guided performance. But sometimes, according to this model, the representations are transformed before they influence performance. According to the Two Modes of Visual Representation hypothesis, the proximal pattern is converted into two representations that are determined by separate neural pathways. One of these representations mediates perceptions and verbal responses, the other mediates motor responses. Finally, according to the EIEIO hypothesis, the proximal pattern is converted into multiple abstract symbolic representations. One of these is formed by general input operations that mediate perceptual impressions and some sensory-guided behaviors. The others are formed by specialized input operations, EIEIOs, which mediate specific sensory-guided skills. The general input operations are the most robust in that they are adapted to operate optimally over the entire range of variability to which the system is exposed. This robustness is gained at the expense of efficiency and accuracy in any given situation. For example, the processing efference-based and light-based information in a well-lit, structured environment might be less efficient than the processing of light-based information alone, but this strategy would protect an organism that is suddenly confronted with a situation in which the light-based information is reduced. In contrast, EIEIOs develop to serve sensory-guided skills optimally in a specific context. These input modules are extremely powerful in that context, but are very vulnerable to failures outside that context.

I originally postulated the existence of EIEIOs to account for highly skilled sensorimotor performance of athletes, pilots, and astronauts. I then realized that they might also apply to more common, highly practiced skills such as grasping, catching, or hitting objects within arm's reach. The ball and cup demonstration is consistent with this possibility. Accordingly, perceptual impressions in that situation are mediated by general input operations that are relatively robust to the elimination of binocular information because redundant monocular information is also processed. In contrast, hitting responses in that situation are mediated by an EIEIO. The results suggest that this particular input module is more dependent upon binocular information. This strategy might have provided the EIEIO with greater efficiency in one common situation, but sacrificed robustness in other situations.

It is one thing to consider that a select few of our species, such as athletes and pilots, develop specialized event input operations to service their extremely high level sensorimotor skills. It is quite another to suggest that we all do it to control ordinary skills such as grasping, catching, and hitting in our everyday lives. An implication of the latter possibility is that the domain of perception with respect to perceptual impressions, and the domain of perception with respect to sensory-guided performance, might be more distinct than we had realized. Consequently, we might have to modify our analytic approaches to these domains. Past analyses of the nature and determinants of perceptual impressions have yielded fundamental principles such as the laws of organization. Do these principles apply to the input operations that underlie sensory-guided performance? The present considerations suggest that this question must be answered by empirical tests rather than by assumptions. The uniqueness of the EIEIO hypothesis is in the heuristic implications for such tests.

After my presentation I was asked to explain how the EIEIO hypothesis differs from other modularity models. I will conclude by answering this question. A salient feature of the EIEIO model is that it includes more than one abstract, symbolic representations of space, only one of which corresponds to perceptual impressions. As illustrated in Fig. 1, other models include that characteristic. Summary comments on this conference provided a historical context for consideration of such models (see Stark, this volume). In light of these comments and my own attempts to trace historical roots, I believe that the EIEIO model has not only a novel name, but also unique heuristic merits that will become clearer when more data are collected. In checking out the five premises that are outlined in my paper, I will be testing ideas for which there are no other tests that I have been able to find. The unprecedented experiments will focus on ways in which practice of a sensory-guided skill can reconfigure the way in which input operations utilize proximal information. Two types of experiments are suggested: one that analyzes existing skills, as was done in the ball and cup demonstration, and one that examines the learning of new sensory-guided skills. The focal questions concern the constants and variables of adaptive input operations that underlie relationships between stimulus information, perceptual impression, and performance. The processes underlying the laws of organization might be examples of processes that are universal and constant across all input operations. But, as noted earlier, the EIEIO hypothesis indicates that such possibilities must be tested rather than assumed.

Given the limited scope of this commentary, I can only paint in broad strokes the kind of tests that are suggested to me by the EIEIO heuristic. The tests that I am planning were greatly influenced by work summarized by Marr (1982). He provided detailed models of lower visual processes at three levels of explanation: (1) computation, (2) representation and algorithm, and (3) hardware (neural) implementation. In contrast, models of higher processes were limited to the computational level and were much less developed. A sharp decline in detail occurred in modeling the transition from a viewer-centered frame of reference (two and one-half-dimensional sketch) to a three-dimensional frame of reference based on the shape itself. Marr stated that an obstacle to more detailed modeling of these higher processes is the difficulty of discovering "what systems and schemes are actually used by humans...at present I see no empirical way of approaching this type of problem. It seems to be much more difficult to design experiments to answer questions at these rather high levels of analysis than at the lower ones...Designing a successful empirical approach to such questions would represent a major breakthrough." Experiments that gave major insights into lower-input operations were often based on dramatic perceptual impressions, such as those created by Julesz's random dot stereograms or by Ullman's rotating cylinder demonstrations. Higher operations, such as those that underlie object recognition and/or localization, are much more difficult to capture with such demonstration because of the variable and complex relationships that exist between stimulus information, perceptual impressions, and performance.

In order to account for these many-to-one relationships, Marr proposed a model that bears directly upon the present considerations. He suggested that a single, two and one-half-dimensional sketch is constructed in order to serve all sensory-guided systems, and that different systems process this abstract, symbolic representation according to different rules to suit different purposes. Since the data employed in testing the nature and determinants of the two and one-half-dimensional sketch were based on perceptual impressions, Marr's model can be interpreted as a Perception Plus Transformation model.

The EIEIO is similar in suggesting distinct input modules for different purposes, but the EIEIO model does not assume common operations for all modules up to the level of a two and one-half-dimensional sketch, or up to any other abstract representation. Instead, the EIEIO model leaves open for testing the possibility that separate input models already diverge at the initial sampling of the proximal pattern, which is defined at the interface between physical information and sensory receptors before abstraction processes begin.

This contrast between models suggests a starting point for testing. My plan is to use displays similar to those that have cast light on processes that yield a two and one-half-dimensional sketch. One such display is Ullman's counterrolling cylinders, which consists of a sequential presentation of a set of frames. Each frame is a random set of dots, and that is how each frame appears when it is presented alone. The relationship between frames, however, is highly structured such that the frames present a screen containing successive orthographic projections of two concentric cylinders that are counterrotating. When the frames are presented at the appropriate rate, observers see counterrotating cylinders. This perceptual impression was Ullman's main response measure. I will modify the display in order to manipulate monocular versus binocular viewing, stereopsis, texture gradients, brightness gradients, and other information about the screen's orientation and distance. I will also add both verbal and motor response as well as more task demands and response measures, such as more detailed reports of perceptual impressions as measured by Epstein and Park (1986), measurements of forced-choice recognition, measurements of viewer-centered surface orientation and distance by means of alignment of an unseen body part with the surface, measurements of object-centered surface orientation by means of comparison with a standard object, and measurements of accommodation and convergence. An initial step will be to replicate the Old McDonald effect in this context and to pursue other discrepancies between perceptual impressions and performance, including recognition and visuomotor coordination. In addition, I will try to create such discrepancies by selectively manipulating sources of information during training session on different tasks.

An important phase will be testing opposing predictions of Perception Plus Transformation models and the EIEIO model. For example, control over separate sources of information will enable precise manipulations of the degree of veridicality of perceptual impressions. Perception Plus Transformation models will be supported whenever recognition or localization responses are related to perceptual impressions by a transformation rule; the EIEIO hypothesis will be supported whenever sensory-guided performance and perceptual impressions vary independently. Finally, the Two Modes of Visual Representation hypothesis will be tested by comparing verbal and motor responses.

The proposed empirical approach that was suggested by the EIEIO heuristic is a hybrid of methods traditionally used to measure perceptual impressions such as the constancies, and methods that have been used to analyze cognitive processes such as stages of processing in pattern recognition. The approach is aimed at two goals; (1) to provide a data base for inferring the systems and schemes that determine perceptual impressions and sensory-guided performance; and (2) to advance spatial instrument technology by enhancing our ability to understand, predict, and control the many-to-one correspondence that often exists between stimulus information, perceptual impressions, and performance.

REFERENCES

Epstein, W.; and Park, K. (1986): Continuous optical transformations do not elicit unique perceptual description. Percept. Psychophys., 40, 365-469.

Marr, D.(1982): Vision, New York: Freeman.

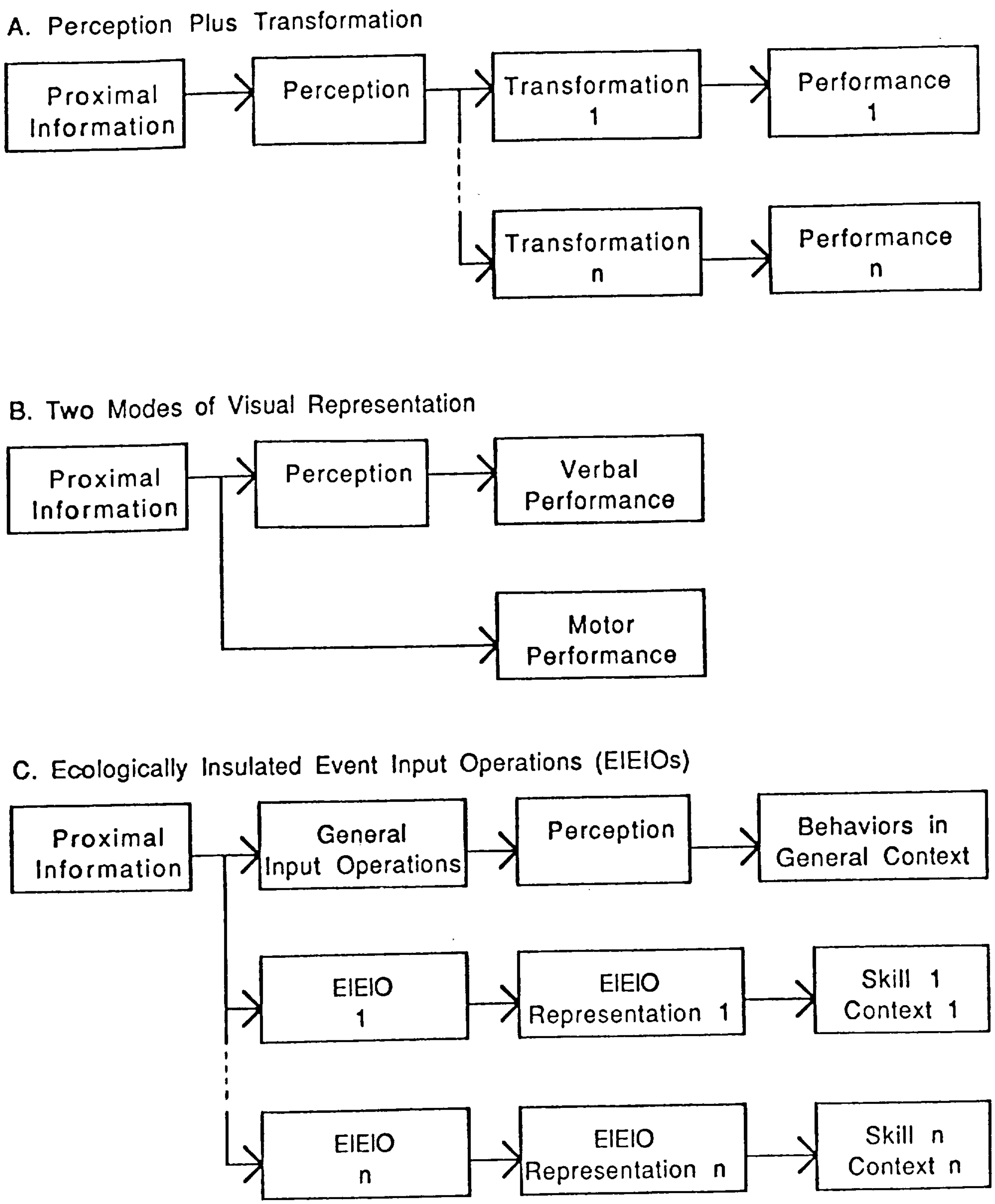

Figure 1.– Hypotheses proposing framework within which to investigate the many-to-one relationship existing between stimulus information, perceptual impressions, and performance.